AD-A065 452 NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA F/G 5/7

NEW METHODOLOGIES FOR AUTOMATED DATA FUSION PROCESSING.(U)

SEP 78 R A DILLARD

UNCLASSIFIED NOSC/TR-364 NL

| OF |

AD
A065452

END
DATE
FILMED
4 --79
DDC

NOSC

(12) LEVEL II
NW

NOSC TR 364

Technical Report 364

NEW METHODOLOGIES FOR AUTOMATED DATA FUSION PROCESSING

RA Dillard

September 1978

Interim Report: June 1977–September 1978

Prepared for
Naval Electronic Systems Command

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

RR GAVAZZI, CAPT, USN
Commander

HL BLOOD
Technical Director

ADMINISTRATIVE INFORMATION

The work reported in this document was performed at NOSC for the Naval Electronic Systems Command project XR01408. Assistance is greatfully acknowledged to Lt. C Coghill, SF Fickas, DC McCall, and Dr. TH Crocker for discussions on many examples and applications used herein.

Released by
RC Kolb, Head
Tactical Command Control Division

Under authority of
JH Maynard, Head
Command Control –
Electronic Warfare Systems
and Technology Department

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

14 NOSC/TR-364

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOSC Technical Report 364 (TR 364)	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 NEW METHODOLOGIES FOR AUTOMATED DATA FUSION PROCESSING		5. TYPE OF REPORT & PERIOD COVERED 9 Interim Report 1 June 1977-30 September 1978
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Robin A. Dillard		8. CONTRACT OR GRANT NUMBER(s) 17 XR0140801
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego, CA 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 16 61153N, XR01408, XR0140801, 824-CC25
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Electronic Systems Command Washington, DC		12. REPORT DATE 11 September 1978
		13. NUMBER OF PAGES 37
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 12 40 p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

DDC RECEIVED MAR 8 1979 B

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Artificial intelligence
Data fusion
Knowledge representation
Natural language processing
Pattern recognition
Production systems

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report initiates investigations into some of the newer technologies to determine if they have useful applications to the problem of automating the fusion of multisource data. The main technology areas discussed are natural language processing (NLP), production rules, and pattern recognition. The theory of possibilities is also considered.

An integrated data fusion system is postulated which would employ a number of different, interacting techniques. Early stages of system processing would require NLP techniques to restructure narrative data into

393 159

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 03 05 037

20. (Continued)

a form useful for automated fusion with other kinds of data. The use of "frames" as knowledge representation structures is considered for this part of the processing. Techniques involved in later processing could include production rules and pattern recognition. Possible applications of these two techniques to data fusion problems are discussed.

OBJECTIVE

The military decision maker is unable to effectively use, in a timely manner, the increasing amount of diverse data available. The purpose of this task is to investigate the newer technologies to determine if they have useful applications to the problem of automating the fusion of multisource data.

RESULTS

1. The special requirements in data fusion for computer understanding of textual material are identified and some current approaches to natural language processing (NLP) are examined for their suitability. It is concluded that a "frames" approach to NLP, with a significant amount of additional enhancement, could probably provide a suitable conceptual structure for representing the narrative information in Navy messages.

2. Production systems are found to have several features attractive for data fusion applications. System organizational aspects such as weighting mechanisms and net structure are examined in an example of an application to platform identification.

3. It appears that there are several possible applications of pattern recognition involving multisource measurements. A few examples are given.

4. After a brief, initial look at possible applications of the theory of possibilities it is concluded that, at several points in fusion processing, fuzzy-set computations are appropriate under certain circumstances.

5. An integrated data fusion system, which would employ a number of different interacting techniques is postulated and a descriptive model of such a system is presented.

RECOMMENDATIONS

The follow-on effort should include a continuation of the investigations of individual techniques and an integration of the more promising techniques into a small-scale experimental model of the postulated data fusion system. Specific steps are described below.

1. Investigate production system organizations possibly suitable for data fusion applications by experimenting with scaled-down sets of rules and data.

2. Study, in more detail, the special problems encountered with textual material in Navy data fusion, e.g., ellipses and the continued need for updating, and investigate ways of adapting text-understanding techniques to meet these special needs. Later, experiment with knowledge representation structures that might be suitable for interfacing restructured textual data with automated fusion processes, by first integrating a text-understanding process with the experimental model of a production system.

3. Continue to build a small experimental model of a data fusion system by integrating other processes with the experimental production system and text-understanding processes. Use this experimental model to find the interactions among these various processes.

CONTENTS

ILLUSTRATIONS

I. INTRODUCTION

The conversion of increased masses of multisource data, now available to Navy commanders, into relevant information from which determinations of threats and of resource capability/availability can be made is difficult. This problem worsens as new sensor systems are developed and communication capabilities are expanded. Also, weapons ranges are increasing; so an at-sea commander needs to know the locations and activities of his own and hostile forces over a region much greater than that covered by his own ship's sensors. In many situations, and especially at sea, the data fusion process of evaluating, integrating, interpreting, and analyzing the data will require an amount of manpower far in excess of that which we can expect will be available; also, the human fusion process cannot always cope with situations that require a reaction time of a few minutes.

The answer to this problem, of course, is to automate data fusion processes wherever possible. There are a number of techniques emerging in newer technologies which appear to be applicable to the automation of data fusion. The purpose of this project is to investigate some of the newer technologies to determine if they have useful applications to the data fusion problem. The main technology areas examined are natural language processing, production rules, pattern recognition, and the theory of possibilities. In this report, the results of these initial investigations are described.

II. AN OVERVIEW OF AUTOMATED DATA FUSION

A. THE SYSTEM CONCEPT

Our study of the problem has led us to the conclusion that automated data fusion will require the integration of many interacting subprocesses. Figure 1 is a simplified functional diagram of a hypothetical integrated system that we believe has the necessary

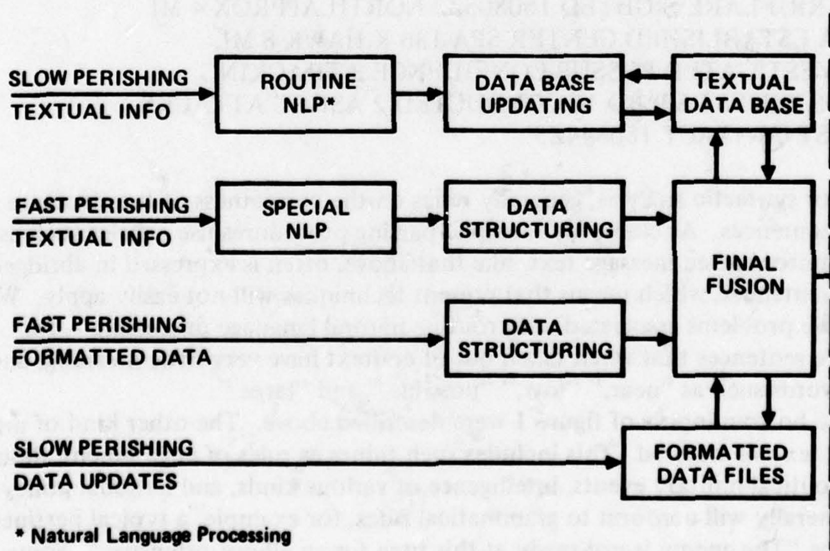

Figure 1. Integrated data fusion processes.

attributes for automating the fusion of the various kinds of data that must be dealt with. Each kind of data has to be processed differently; and, in order for it to be automatically fused with other kinds of data, it has to be restructured into a form that is compatible with the various final fusion processes that it will be involved in.

B. DATA TYPES

One category of slow perishing information is that which can be stored in files requiring only occasional updating. For example, for each kind of platform, both own force and hostile, there are lists of weapons and lists of sensors; and for each weapon or sensor there is a list of capabilities and characteristics. Also, there are deviations in weapons and sensors among platforms of the same class, and these must be listed. There must be data on expendables such as fuel, water, food and medicine. In general, there are many kinds of data that need to be updated, at the most, only a few times a day. These are examples of slow perishing formatted data.

A greater problem exists with fast perishing data. Tactical data systems such as NTDS provide data that often must be dealt with quickly. Also, there are formatted messages arriving via a number of communication links. Examples of formatted lines in messages are the following:

AREA/4200N6/16500E2/100NM

CREW/13/TC-LCDR CLARK/PC-LT ANDERSON/DBD-ENS FISHER

ELLIP/5230S0/03180W2/130T/80NM/60NM/12600SQNM

In addition to the fast perishing formatted data, we have a lot of fast perishing unformatted data to deal with. The (fictitious) message text below illustrates some of the special problems that the Navy will have with natural language processing.

NARR/FLARE SIGHTED 180805Z2 NORTH.APPROX 4 MI
SPA ESTABLISHED.CENTER SPA 136 K HAWK 8 MI.
INVESTIGATED POSSUB.CONFIDENCE 3 TRACKING
NORTHWEST SPEED 16.CONDUCTED 2 ASROC ATTACKS.
LOST CONTACT 180844Z5.

Parsing, or syntactic analysis, generally relies on the correctness of the structure and grammar of the sentences. As clues, for example, parsing procedures use articles such as "a" and "the." But unformatted message text, like that above, often is expressed in abridged and incomplete sentences, which means that present techniques will not easily apply. We also have all of the problems associated with routine natural language processing – for example, we have sentences that when taken out of context have very little meaning, and we have imprecise words such as "near," "low," "possible," and "large."

The three bottom inputs of figure 1 were described above. The other kind of input is slow perishing textual material. This includes such things as rules of engagement, descriptions of world political-military events, intelligence of various kinds, and national policy. This material generally will conform to grammatical rules; for example, a typical pertinent sentence might be "The enemy is not ready at this time for an all-out offensive." Some of

the textual material, such as pacts and treaties, will have a highly organized paragraph structure. Two examples of this are shown below.

Example 1

Article 1. It is forbidden:

1. To . . .
2. To . . .
3. To . . .

Example 2

ARTICLE 1

For the purpose of this Agreement, the following definitions shall apply:

1. "Ship" means:
   a. . . .
   b. . . .
2. "Aircraft" means . . .
3. . . .

C. VOIDS IN THE SYSTEM

Figure 2 is an expanded illustration of the top row of processing shown in figure 1. The text-understanding system would be a subsystem of the postulated total integrated system. It appears that several natural language processing approaches now being developed by researchers (Section III.A discusses some of these) can, after several years of further development, be adapted to the problem of converting slowly perishing textual material into a useful data base. So it probably will not be necessary to do research in that particular area except to determine if the structured data will be in a form, or can be restructured into a form, that is suitable for automated fusion with other kinds of data. However, the problem of updating a textual data base is not receiving much attention from the research community. This project will have to address that problem to some degree in next year's effort. The input to the updating box would be coded, conceptually structured information. The elements of information in the data base would be conceptually bound to other elements within the category and to elements in other categories.

In general, systems that query data bases have human users, and the newer systems being developed accept natural-like language inquiries. In our situation, the queries would be automated requests for specific information from an automated fusion processor, and we will have to investigate possible ways of doing this, although this automated query problem is not an immediate or a major effort here.

Our major problem in natural language processing will be with fast perishing data that are not formatted. As was illustrated in the earlier example, it requires a special kind of processing because of the abridged and incomplete sentences.

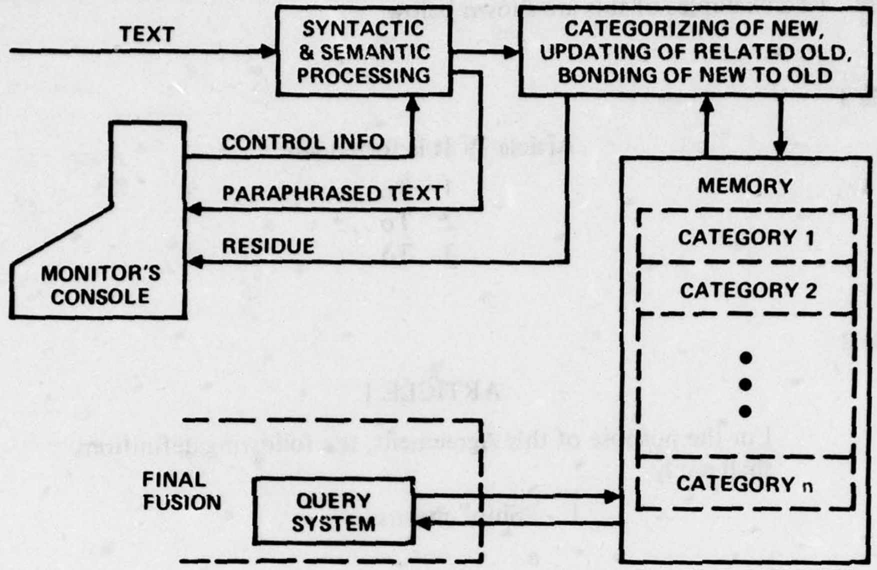

Figure 2. A text understanding system.

The contents of the final fusion box in figure 1 are still very nebulous, but we do recognize that final fusion must involve a number of different kinds of subprocesses that must interface with each other and with structured data derived from natural language sources. In connection with this final fusion box, the use of production rules and pattern recognition will be discussed in Section III. The output of the final fusion box could be, for example, identifications of ships and other platforms, determinations of enemy capabilities, and predictions of enemy actions.

Many of the actions and events that must be dealt with in data fusion involve movements of objects on or near the Earth's surface. The process of automatically associating these events often will require the kinds of computer calculations now augmenting human data fusion plus new algorithms, based on classical mathematics, that will replace the human function of plotting and measuring and of concluding geometrical facts from the plots. These analytical computations, some of which do not yet exist in the form of computer programs, must be interfaced with or, in some cases, interspersed or imbedded in the artificial intelligence processes. These necessary analytical processes are mostly disregarded in this report, but it should be recognized that their implementation would require a sizable effort in the development of an automated data fusion system.

III. TECHNOLOGIES AND THEIR APPLICATIONS

Possible applications to data fusion are discussed in this chapter for three technology areas: natural language processing, production rules, and pattern recognition. Although the three kinds of techniques are discussed here in their pure forms, in practice they likely would be intertwined not only with conventional statistical techniques and data base management techniques, but also with each other and other artificial intelligence techniques.

Natural language processing will be essential to automated data fusion, both for interfacing the fusion system with the system users and for converting narrative data into a form usable by automated fusion processes. It will be seen in Section A of this chapter that natural language processing techniques, when fully developed, can be used not only to restructure textual data, but also to fuse narrative data with formatted data which accompanies it or with narrative text on the same topic from another source. This partially fused data would then be used by other automated processes.

Production systems and pattern recognition are two of a number of techniques that might be useful in the box labeled "final fusion" in figure 1. Production systems, which are discussed in Section B, can substitute for human reasoning processes by representing a wide range of world knowledge in the form of premise → conclusion rules. Computation time and memory requirements are serious problems to be encountered, but with proper system organization these might be held to reasonable levels. Pattern recognition would be limited to small, well structured problems such as recognizing maneuvers and course variations given track data and, for example, locations of sensors, submarines, or weather that a ship should avoid. Other examples of applications of pattern recognition are given in Section C.

A. NATURAL LANGUAGE PROCESSING (NLP)

Background

The data inputs to the fusion process were described in Chapter II, where they were divided into four general categories. NLP techniques are needed to deal with data in two of these categories: slow perishing textual information, which is usually in the form of a well-written document, and fast perishing textual information, which is typically the comments section of a tactical message. As explained in Chapter II, our objective does not include developing NLP techniques, but we must investigate the NLP techniques currently under development to find those that show promise of providing a suitable interface with automated fusion processes. In our investigation of NLP techniques, we hope to determine: (1) the extent to which NLP techniques will be satisfactory or inadequate for this application, (2) what must be provided by the future user of the fusion system (e.g., special vocabulary, unusual grammatical characteristics, facts about military equipment and operations, etc.) in order to employ an NLP method, and (3) what the conceptual structure of the processed textual data will be, relevant to compatibility with automated fusion processes.

This section summarizes the expected problem to be encountered with NL data and describes some of the approaches to coping with these problems. The problems relating to textual context are discussed first.

Context and Frames

Natural language processing would be relatively simple if sentences were self contained packets of information, but a sentence taken out of context often carries little information. For example, consider some of the possible meanings of the report: "WILL TAKE THEM IF LIGHT."

1. The helicopter will carry the wounded to the carrier if the gunfire is light.

2. If the tank landing ship reaches the treacherous shoreline before dark, the troops will capture the terrorists.

3. The cruiser personnel will have inoculations if the package of serum is light enough for transport in the heavily loaded helicopter.

This is an extreme and unlikely example, but it illustrates two of the more common contextual problems: (1) many words have multiple meanings, and (2) pronouns substitute for nouns. Even an expansion of this five-word line into a correct sentence, "We will take them if it is light." is of negligible help. First we need to know who the actors are, what their mission or goal is, and what relevant events have previously been reported.

While some of the information needed to understand a sentence is contained in other sentences, in many cases much of the information needed is not contained anywhere in the textual material but must be inferred by the reader or listener. This is possible because of the reader's experience and so-called common sense.

Because of these problems with context, most researchers in the area of text understanding by computers use "frames" in their approach to knowledge representation. A frame is a data-structure for representing a situation, and can be thought of as a network of nodes and relations (refs. 1–3; also ref. 4).

In his examples of how to represent a situation with a frame, Minsky (refs. 1, 2) considers such situations as: being in a certain living room; going to a child's birthday party; looking at a cube, a table, a chair partly hidden by a table, a flowing river, and a car generator; visualizing the workings of a car generator from a mechanical viewpoint; visualizing the workings from an electrical viewpoint; and using a piggy bank.

Data fusion applications of a frame approach would include situations that involve a usual sequence of events, such as a refueling procedure, a missile attack (from targeting solutions, launching, mid-course guidance signaling, etc., to damage assessment), a particular kind of training exercise, an infrared flare ejection (which should bring to mind an aircraft threatened by a heat seeking missile) and a submarine rising to periscope depth (the detection of a periscope suggests the presence of a submarine just below and the possibility of a transmission). Relatively static situations which possibly could be represented by frames are: "looking" at a land mass (reference to a unique landmark, for example, could call up other pertinent information about an otherwise unidentified area), and "looking" at a ship's superstructure (enabling reasoning of the type "if its superstructure was badly damaged, its surface-search radar is probably inoperative"). Intelligence reports of various kinds might also be representable by frames. A report that "country x plans to achieve domination of countries y and z by aggressive political efforts and by a threatening show of naval strength in the Gulf of . . ." should invoke a frame which recognizes that political efforts and shows of strength are methods of achieving domination, and that there is an increased

---

1. Minsky M, A Framework for Representing Knowledge, in the Psychology of Computer Vision, ed. PH Winston, p 211-277, McGraw-Hill, 1975.
2. Minsky M, Minsky's Frame System Theory, in Proceedings of the Conference on Theoretical Issues in NLP, p 104-116, Cambridge Mass, June 1975.
3. Winston PH, Artificial Intelligence, Addison-Wesley, 1977.
4. Hewitt C, Stereotypes as an ACTOR Approach Towards Solving the Problem of Procedural Attachment in FRAME Theories, in Proceedings of the Conference on Theoretical Issues in NLP, p 94-103, Cambridge Mass, June 1975.

expectation of country-y's naval units moving, in response, to that area. (We would desire, further, that the show-of-force information be tagged as important, while other information simply be held available, in its restructured form, for later explanation traces requested by a user.)

Minsky (ref. 2) explains that the essence of his frame theory is: "When one encounters a new situation (or makes a substantial change in one's view of a problem), one selects from memory a structure called a frame. This is a remembered framework to be adapted to fit reality by changing details as necessary." (There is not full agreement on this concept and other aspects of Minsky's frame theory; ref. 5 sites some problems.) The top levels, or layers, of a frame network represent things that are always true about the situation, while the lower levels have "slots," called *terminals*, that must be filled with data derived from the textual material about the situation. Each terminal can specify conditions on the data to be assigned to it.

Generally, the representation of narrative text would require collections of related frames linked together into "frame systems." For visual scene analysis, for example, different frames of a system would represent the scene from different aspects. For nonvisual frames, the links between frames can represent changes of emphasis and attention, or cause-effect relations. (Recall the three views of the car generator in the earlier examples.) The different frames in a frame system share the same terminals. Minsky envisions (ref. 1, 2) that "a great collection of frame systems is stored in permanent memory, and one of them is evoked when evidence and expectation make it plausible that the scene in view will fit it." He proposes that "if a chosen frame does not fit well enough, and if no better one is easily found, and if the matter is important enough, then an adaptation of the best one so far discovered will be constructed and remembered for future use."

The various frame systems are linked together by an "information retrieval network," which participates in the selection of the frame best-suited for representing a situation (refs. 1, 2). The interframe structures also can store additional contextual knowledge useful in understanding textual material about a situation. It is not at all clear how an information retrieval network would operate in data fusion applications, but its existence should help to provide the needed flexibility of representation.

Associated with the use of frames in our postulated data fusion system is an additional complexity not shown in figure 1. The structuring of textual information and of formatted data were shown as separate processes in figure 1. When the textual data is a comments section on an otherwise formatted message, the two kinds of data really should be processed together. Figure 3 outlines a procedure for handling messages of this type. The formatted data would play a major part in the selection of frame types and the information from both kinds of data would fill the frames.

It is further conceivable that the filling of a frame would be resumed later as a result of fusion processes that involve the original frame, or that a new frame would replace the original. For example, a later classification of a platform already partly described by a frame would allow use of contextual information about the platform, such as an explanation of intent based on capabilities and behavior. Unless new NL data is received or the original NL data is reprocessed based on new information, however, the use of frames in a later stage of data fusion could not be called NLP.

---

5. Feldman J, Bad-Mouthing Frames, in Proceedings of the Conference on Theoretical Issues in NLP, Cambridge Mass, p 92-93, June 1975.

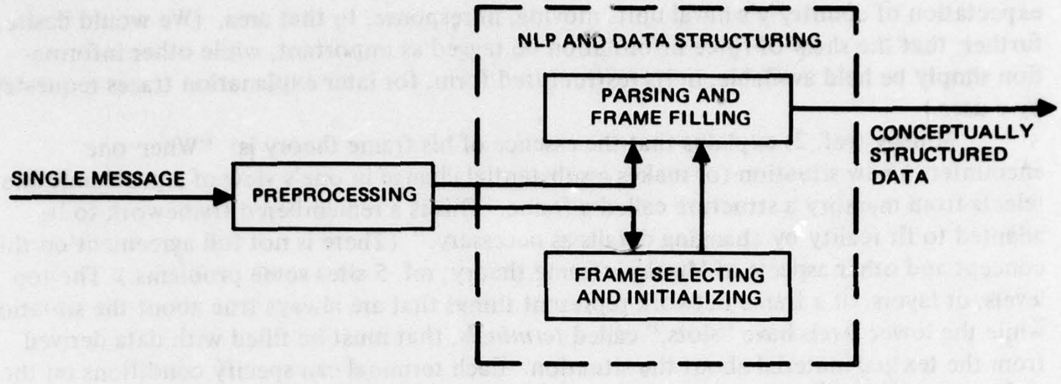

Figure 3. Processing a mixture of data types. The data structuring of the natural language portion of a message should be determined in part by the formatted portion.

Scripts, Plans, Goals, and Themes

Schank and Abelson (refs. 6, 7) have proposed and experimented with a text understanding method that uses specialized versions of frames. Underlying their frames are basic constructions called "conceptualizations," which represent the meanings of sentences. The structures of these conceptualizations must conform to "Conceptual Dependency Theory," which has the basic axiom (ref. 7), "For any two sentences that are identical in *meaning*, regardless of language, there should be only one representation." They postulate two kinds of conceptualizations. An active conceptualization has the form: Actor Action Object Direction (Instrument). A stative conceptualization has the form: Object (is in) State (with Value). Conceptualizations, which are further described in references 8 and 9, involve a number of primitive acts. Among those acts that are applicable to data fusion are ATRANS (the transfer of an abstract relationship such as possession, ownership or control), PTRANS (the transfer of the physical location of an object), PROPEL (the application of force to an object), MTRANS (the transfer of mental information), and MBUILD (the construction of new information from old information).

While Conceptual Dependency representation was designed to handle single thoughts or sentences, causal chains are needed to handle the connections among the sentences and thoughts in a text. In reference 7, the authors point out that a simple causal syntax exists in natural thought, but this syntax can be violated in natural language expression (and they give the example "John cried because Mary said she loved Bill."). In regard to this problem, we mentioned earlier in this section that the reader or listener often must infer the information needed to understand a sentence, in this case to link one sentence or thought with

6. Schank RC, and Abelson RP, Scripts Plans and Knowledge, Advanced Papers for the Proceedings of the Fourth International Joint Conference on Artificial Intelligence, p 151-157, Tbilisi USSR, September 1975.
7. Schank RC, and Abelson RP, Scripts Plans Goals and Understanding; An Inquiry into Human Knowledge Structures, Lawrence Erlbaum Associates, 1977.
8. Schank RC, Identification of Conceptualizations Underlying Natural Language, in Computer Models of Thought and Language, ed. RC Schank and KM Colby, p 187-247, WH Freeman and Co, 1973.
9. Schank RC, The Primitive ACTs of Conceptual Dependency, in Proceedings of the Conference on Theoretical Issues in NLP, p 34-37, Cambridge Mass, June 1975.

another. As a mechanism for dealing with this problem in stereotyped situations, Schank and Abelson (refs. 6, 7) introduce the notion of a SCRIPT, which is defined (ref. 10) as "a predetermined causal chain of conceptualizations that describe the normal sequence of things in a familiar situation." In the fusion of tactical data, for example, a script for a routine surveillance mission would be useful for understanding messages received about mission activities. A script, which is a special version of a frame, is a structure made up of slots and of requirements on what can fill those slots. It describes an appropriate sequence of events in a particular context.

Schank and Abelson (ref. 7) also introduce a more general version of a frame, a PLAN, which is a mechanism used to describe actions in new or unexpected situations (in our use, perhaps a military operation or an order of battle). They explain that "there is a fine line between the point where scripts leave off and plans begin" and in fact allow a plan to call into use a script when appropriate for reaching a subgoal. A plan includes a series of actions to reach a goal, and the method of realizing a goal usually involves a chain of instrumental goals.

While a plan can be explained only in the light of the goal or goals that generate it, goals are not usually explicitly stated in the text but must be inferred from a "theme." Reference 7 names three categories of themes: role, interpersonal, and life themes. The role themes considered there are societal roles which can be referenced by particular English words, such as waiter, president, or psychiatrist. In our application of NLP, the role theme of a platform or agency would be determined by the purpose or mission for which it was designed. For example, "roles" of platforms would be referenced by the labels oiler, attack submarine, minesweeping boat, aircraft carrier, etc. Agency roles are different in that they tend to be unique, e.g., the Naval Ocean Systems Center serves as an RDT&E Center for command control, communications, ocean surveillance, etc. Each role involves many functions or responsibilities, and the pertinent details must be incorporated into the text understanding system. An aircraft carrier, for example, has the obvious function of launching and landing particular airplanes, but also has surveillance, weapons, tactical support, command, control and communications capabilities and responsibilities.

The "interpersonal theme" is less applicable here, but an example might be "The dictator of country x is angry about the withdrawal of military support by country y." A "life theme" might be "Industry is essential to the existence of country z, and they must buy most of their oil from other countries."

A text understanding system, then, should recognize when a goal exists for an actor, in order to understand his (or its or their) actions based on that goal; and unless the goal is stated in the text, it must be generated by the system by using "theme" knowledge about the actor.

Other NLP Techniques

During this initial effort in data fusion we have investigated many of the current NLP techniques, although only a few in depth. All of the NLP schemes are incomplete in many respects, and it is difficult to project their capabilities, if or when fully developed, in

10. Schank RC, Using Knowledge to Understand, in Proceedings of the Conference on Theoretical Issues in NLP, p 117-121, Cambridge Mass, June 1975.

processing Navy textual material. An excellent summary of work in NLP is given by Raphael (ref. 11). Rather than include a summary here, we will simply comment below on recent work of special interest.

Reference 12 describes the system concept of a system that will analyze incoming textual reports of events and, from them, synthesize "event records" (i.e., extract relevant information and store it in a data base record). The technique employs framelike "event templates" for representing knowledge. Because the types of textual reports they consider (the work was sponsored by the Air Force) are very similar to some of ours, their approach deserves attention in our future investigations. Reference 12 also discusses several other current approaches to knowledge representation.

A scheme for using frames in the comprehension of simple narration is described by Charniak in reference 13. The technique is similar to the independently developed "script" described in the previous section, but Charniak structures his frames in a particular way. For example, he permits frame statements in one frame to be shared with another, not physically but by using an identity pointer in one of the frames. Charniak's interest is in the construction of a computer program which will answer questions about simple narration, but some of his ideas on frame organization should be seriously considered in adapting an NLP technique based on frames to the structuring of a textual data base for automated data fusion.

One intriguing technique still in an early state of development is a meaning representation language for natural languages called PRUF (Possibilistic Relational Universal Fuzzy), described by Zadeh in references 14 and 15. The logic underlying PRUF is a fuzzy logic in which truth values are linguistic. PRUF serves as a foundation for "approximate reasoning," a process by which a possibly imprecise conclusion is deduced from a collection of imprecise premises (ref. 16). As an example of PRUF, the report "MERCHANT NEAR MINED AREA EXPLODED." translates in PRUF to the expression:

$$\text{EXPLODED}\left[\text{Subject} = \text{merchant-ship};\ \Pi_{\text{Location}} = {}_{\text{Site 1}}\text{NEAR}[\text{Site2} = \text{mined-area}]\right]$$

where Π_{Location} is a "possibility distribution." (The concepts of fuzzy sets and possibility distributions are too involved to describe here, but are discussed by many authors in recent literature.) While the prospect of a meaning representation language based on fuzzy set theory is promising for dealing with imprecise statements, mechanisms for handling contextual problems seem to be lacking in the presently envisioned PRUF, so its applicability to NLP for data fusion is uncertain.

---

11. Raphael B, The Thinking Computer, WH Freeman and Co, 1976.
12. Silva G, and Montgomery CA, Knowledge Representation for Automated Understanding of Natural Language Discourse: Computers and the Humanities, vol 11, p 223-234, Pergamon Press, 1978.
13. Charniak E, Organization and Inference in a Frame-Like System of Common Sense Knowledge, in Proceedings of the Conference on Theoretical Issues in NLP, p 42-51, Cambridge Mass, June 1975.
14. Electr Res Lab, Univ of Calif Berkeley, Unclassified Memorandum ERL-M77/61, Subject: PRUF – A Meaning Representation Language for Natural Languages, by Zadeh LA, 30 August 1977.
15. Zadeh LA, PRUF and Its Application to Inference from Fuzzy Propositions, vol 2, p 1359-1360 of Proceedings of New Orleans IEEE Conference on Decision and Control, IEEE publication, 1977.
16. Zadeh LA, A Theory of Approximate Reasoning (AR), Memorandum UCB/ERL M77/58, Electr Res Lab, Univ of Calif, Berkeley, 30 Aug 1977.

Interfacing NLP with Fusion

In order to make sense out of textual material, a text-understanding system must have a tremendous amount of world knowledge stored in a suitable conceptual structure. If the text-understanding system were to be a subsystem of an automated fusion system, much of its required knowledge would be the same as that which we would expect to be used by the fusion processes. In this and several other respects, the processing of NL data in a data fusion system could be considered as an early stage of fusion, and not just a generator of inputs. If textual material from several different sources on the same topic are combined and processed as a single story, then we must certainly call this processing a kind of fusion. Also, the use of formatted data in selecting and filling in frames for messages containing formatted and narrative text (fig. 3) would be a form of fusion. Still, it is convenient to treat the two as separate processes that must be appropriately interfaced, while recognizing that some fusion is involved in NLP and some NLP is involved in fusion.

B. PRODUCTION SYSTEMS

Background (references 3, 17–21)

In some areas of human decision making, the reasoning processes can be modelled by rule-based systems. A rule, known in these applications as a "production rule" or a "production," is generally of the form

$$\text{If } F_1 \text{ and } F_2 \text{ and} \ldots \text{and } F_n \text{ then C}$$

or, equivalently,

$$F_1 \,\&\, F_2 \,\&\ldots\&\, F_n \rightarrow C$$

where F_i is a fact, an event, a situation, a string of symbols, or a cause, and C is a conclusion or hypothesized conclusion, an action to be performed, or an effect. Some of the rules in a production system represent the knowledge of trained experts, and others provide system organization.

In addition to an organized set of rules, a production system must have a data base consisting, typically, of gathered pieces of evidence which might be relevant to the condition

17. Davis R, BG Buchanan, and EH Shortliffe, Production Rules as a Representation for a Knowledge-Based Consultation System, Stanford AI Lab Memo AIM-266, Computer Science Dept Rept STAN-CS-75-519, Oct 1975.
18. Davis, R and King J, An Overview of Production Systems, Stanford AI Lab Memo AIR-270, Computer Science Dept Rept STAN-CS-75-524, Oct 1975.
19. Duda RO, Hart PE, and Nilsson NJ, Subjective Bayesian Methods for Rule-Based Inference Systems, SRI-AI Center Tech Note 124, Jan 1976.
20. Hayes-Roth F, "Knowledge representation, organization, and control in large-scale pattern-based understanding systems," Conf Record, Joint Workshop on Pattern Recog and Artif Intel, p 66–73, 1976.
21. Shortliffe EH, Computer-Based Medical Consultations: MYCIN, American Elsevier, 1976.

in the left side of some rule. System organization is provided by several kinds of control mechanisms. An evaluation mechanism is needed to evaluate the left side of a rule based on the evidence in the data base. A rule-selection mechanism determines the order of rule access. It is desirable to have a mechanism for augmenting and modifying the system. A production system also needs direction and weighting mechanisms, which are described further below.

Figure 4 is an illustration of the net structure of a very simple production system. The AND arcs denote single productions (where multiple conditions must be satisfied for the conclusion to follow), and OR inputs are separate productions. The "direction" mechanism of a production system relates to reasoning processes, where inferring and deducing new information from evidence can be considered opposite in direction from hypothesizing and then testing the hypothesis. One type of system direction is forward running; these systems start with input data and proceed up to conclusions. Backward running, or top down, start with hypothesized conclusions that are selectively generated and proceed to see if they are supported by the data base. Some systems use an ad hoc combination of up and down directions.

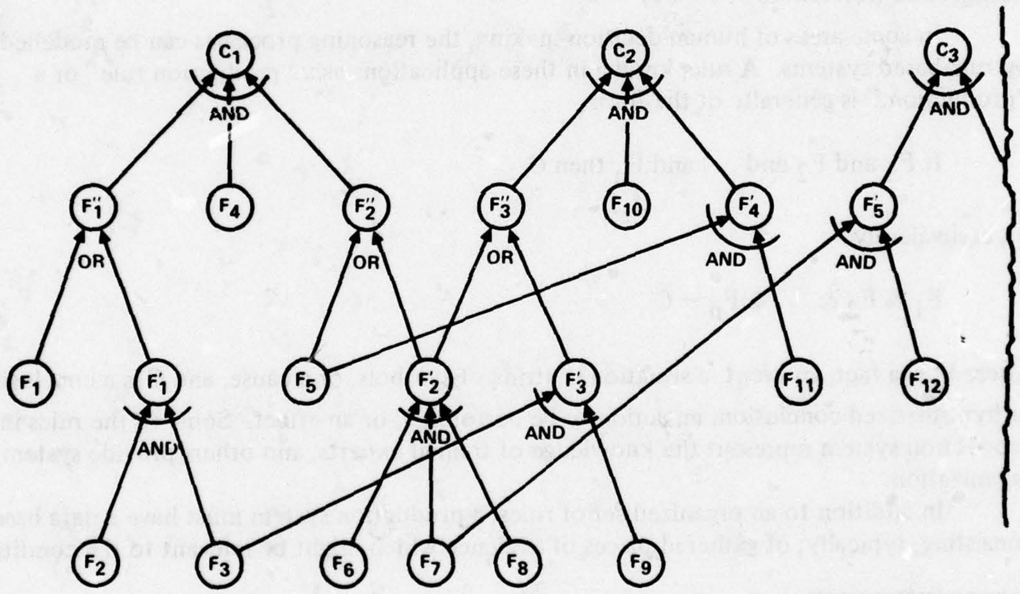

Figure 4. Trees of conclusion in a production system. Symbols ' and " denote intermediate conclusions which are deduced facts used in later productions.

When using a production, there is often associated with each F_i in the premise a quantity, known as a "certainty factor," which indicates the likelihood that F_i is true based on the input data. Also, for most production rules, the premise leads to the conclusion with, say, an 80% or 90% probability, instead of absolutely. Similarly, there may be a significant probability that the conclusion is true even when the premise is not satisfied. Measures of the latter two likelihoods are known as "strengths," "attenuation factors," or "certainty factors." All of these quantities can be used in estimating the certainty factor of a conclusion.

Many conclusions are intermediate conclusions that are then treated as facts for future productions. "Weighting" is a term that refers to these quantities and their propagation through the net. Weighting can be used to determine the reliability of final conclusions and also to reduce the number of computations through the pruning of unlikely hypotheses.

If the statistics of the process are known sufficiently, Bayesian weighting can be used. Bayesian weighting is discussed in detail in the appendix. A more common method of weighting is to use ad hoc scoring functions. When the conditions F_i or the evidence about them cannot be considered independent, fuzzy set theory can be applied. For example, as pointed out in reference 19, the fuzzy set computations $P(F_1, \ldots, F_n) = \min P(F_i)$ can be used at AND nodes. A weighting technique that uses judgmental measures of belief and disbelief in a hypothesis will be discussed in a later section and described in more detail in the appendix.

Production Systems in Data Fusion

The use of production rules is a possible tool in automating data fusion. The technique could be employed in several different applications in the final fusion box of figure 1, along with other kinds of techniques. For example, a production system could serve as an alerting system for various kinds of critical situations. An application to platform identification is discussed in detail in the next section.

An advantage of a production system is that it can be designed to provide high user confidence. The user can read the lists of rules and can question any conclusion (in a sophisticated implementation), and the system can present to him the facts and logic leading to the conclusion. If he disagrees he can change the rules; with an appropriate mechanism for modification and augmentation, modular pieces of knowledge in the form of production rules can be added or changed without difficulty. In automated fusion applications, these system attributes are especially important. A user is unlikely to accept the system's conclusion if he does not understand the logic behind it or previous conclusions. And he must be able to easily correct or refine the system and to incorporate new knowledge into the system when changes occur in hostile force procedures or equipment.

Aside from the obviously difficult task of acquiring rules, there are several special problems that will be encountered in applying production systems to fusion problems. At the system front-end there is the problem of evaluating the left side of a rule based on the conceptually structured data obtained through the processing of natural language reports. The difficulty of this problem was noted earlier in a discussion of figure 1. In platform identification applications, much of the data will be inaccurate or even totally wrong because of deception or human error. If the density of unknown platforms is high, many hypotheses must be considered and multiple conclusions are needed. Moreover, conclusions will often have to be updated because of the continual arrival of new data. There are also the geometrical problems of track association to be solved, but these are inherent in any approach to data fusion, manual or automatic. Probably the greatest problem with production systems or any automated system is that there are innumerable nonroutine situations which could occur. While a human might be able to fuse the data in an intelligent way in many of these situations, he probably would not be able to foresee the possibility of these situations in time to incorporate the necessary knowledge into an automated system.

Application of Production Rules to Platform Identification

A decision about the identity of a ship or other platform is generally based on accumulated evidence, where each observed feature or small bit of evidence contributes to the reaching of a conclusion about its class or identity. Because of possible enemy deception and occasional very bad errors or misinformation, it is usually unwise to allow a single piece of evidence, via a rule, to reject or accept a hypothesis about the identity. The process of reaching conclusions based on accumulated evidence in a production system is handled mainly by the weighting mechanism. First we will look at the special problems involved in propagating weights through the OR-nodes in a production system net (fig. 4) for platform identification. The two production rules

- If a platform uses a radar erratically
  then it is probably not a merchant.
- If a platform maneuvers
  then it is probably not a merchant.

involve two different kinds of events. In this case, the weighting mechanism should operate on the weights (the uncertainty of the data and the attenuation factors or strengths of the productions) in such a way that the certainty of the conclusion (that the platform is not a merchant) is considerably greater if both premises are satisfied than if only one is. In other situations, the premises may not always be independent. Consider an OR-node that includes these rules.

- If a platform dodges known sensors
  then it is probably not a merchant.
- If a platform follows bad weather
  then it is probably not a merchant.
- If a platform changes course when (our) radar is turned on
  then it is probably not a merchant.
- If a platform maneuvers
  then it is probably not a merchant.

If two or more of these different course variations are noted at distinctly different times and each triggers a different production, then the premises can be treated as independent. However, it could be difficult to distinguish by a platform's action which premise is satisfied. The evaluation mechanism should allow the same event to trigger two or more productions (based on different interpretations of the same event), but never with data certainty factors (assuming they are probabilities) that sum to a value greater than the probability that the course deviation occurred. (This is also true for productions not meeting at OR-nodes; e.g., when independent classifications of signals from the same emitter can lead to different conclusions.) If it is easily possible for two or more independent interpretations of the same action or event to slip into the data base then a safe weighting procedure at an OR-node would be, for example, to allow only the production with the maximum resultant weight to pass its weight on to the node.

Next, consider a case where we are completely safe in specifying the identity of a platform based on, say, three particular features or actions which are noted with high certainty. Under some kinds of system organization, the resulting certainty factor of the conclusion will be relatively small if there are numerous other possible attributes that can

also contribute to identification, even though the conclusion should follow with certainty. Further, the observations of only two of these three along with an additional two others could give a larger certainty factor, while with this combination the conclusion should follow with less certainty. At this point it is advisable to consider the net structure and the ways in which rules can be combined.

The very simplest production system for identifying platforms would have only a few layers in its net structure and would use a weighting mechanism of the accumulated evidence type, with ad hoc scoring. An illustration of this type of structure is given in figure 5.

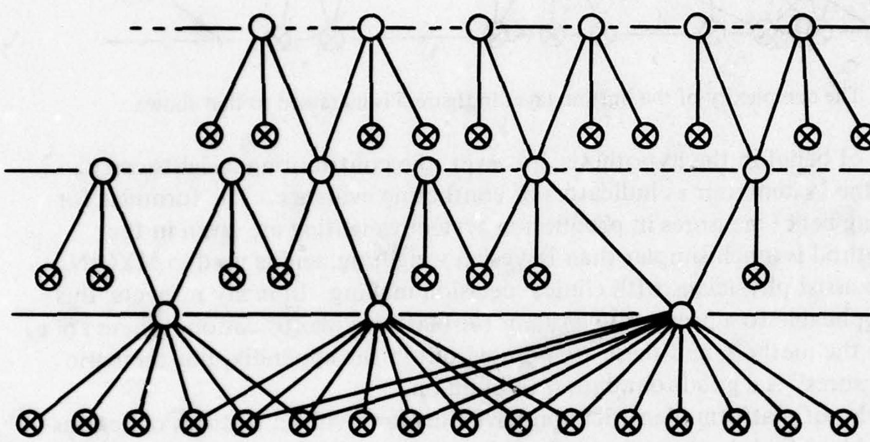

Figure 5. Simple net structure for accumulated-evidence types of weighting mechanisms.

In order to avoid the problem of reaching a conclusion with a small certainty factor even though the several pieces of evidence point unquestionably to that conclusion, we must increase the complexity of the rules, taking into account dependency among conditions. For example, the bottom layer of the simple structure shown above would expand to one like that shown in figure 6. Increasing system complexity in this manner does not make the problem of formulating a weighting mechanism any easier, but it gives us a structure in which we can construct a reasonable one.

While an operator might reach a negative conclusion such as the conclusion "then it is probably not a merchant" given in the earlier examples of rules, a rule when implemented in a production system more normally would use the positive form, "then it is a merchant." When ad hoc scoring is used, evidence against the platform being a merchant then would cause a subtraction from its previous score (or, equivalently, an addition to non-merchant scores). Although a hypothesis is generally stated in its positive form for testing purposes, there is one argument for embodying both the positive hypothesis and its opposite in applications of this type. Often much of the evidence will be contradictory, some supporting one hypothesis and some supporting its opposite. By carrying opposite hypotheses the system can also provide information about the presence of strong, contradictory evidence.

Reference 21 describes a method of weighting which generates weights for opposite hypotheses in the sense that it uses separate measures of belief and disbelief in a hypothesis. These measures are combined into a certainty factor, a number between −1 and +1 that

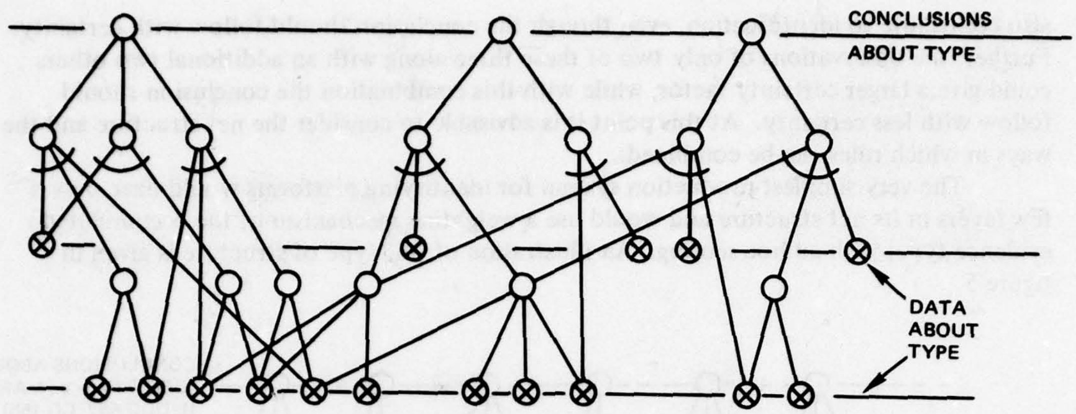

Figure 6. The complexity of the bottom layer in figure 5 is increased to that shown.

reflects the degree of belief in the hypothesis; however, the contributing weights can be made available to the system user as indicators of conflicting evidence. The formulas for this method of using belief measures in production system weighting are given in the appendix. The method is much simpler than Bayesian weighting, and is used in MYCIN, a system designed to assist physicians with clinical decision making. In many respects, this method appears applicable to a production system for platform identification. There are a few problems with the method, and these are pointed out in the appendix, but the basic idea of "belief measures" is a good foundation to build on.

Our examples of platform identification have thus far involved features or actions of the platform to be identified, but there are other kinds of rules that can be used for identification. Consider a platform (call it ship k) that has just been detected and its position fixed. Rules such as the following use a list of platforms (and position data) whose locations were known earlier.

- If the maximum velocity of ship j (from this list) is less than that
  required to reach the location of ship j
  then ship k is not ship j.
- If ship j could not reach the location of ship k by any course without
  being detected enroute
  and if it was not detected
  then ship k is not ship j.
- If ship j could have reached the location of ship k
  and if no other ship could have reached the location of ship k
  without being detected enroute
  then ship k is ship j.

Also, we have conveniently disregarded production rules that would operate in a top-down manner in the net, such as

- If the platform is (the name of a cruiser)
  then the platform is a cruiser.

These rules present no special problems, but are inconvenient to include in figures of system organization.

C. PATTERN RECOGNITION

Background (references 22–25)

Most of the pattern recognition problems that occur in military situations occur before the data fusion stage. Examples of these are the classification of radar signatures, the classification and fingerprinting of intercepted signals, and multitarget radar tracking. Also, for automated fusion, character recognition is needed in a text-reading system. In this task we need to consider applications such as those only as they affect the characteristics of input data to be fused. The main types of pattern recognition that are being considered in this task are those involving patterns derived from multisource inputs, such as platform identification and situation recognition.

Figure 7 illustrates the two main functions of a pattern recognition system. The vector consisting of the feature measurements $x_1, x_2, \ldots, x_n$ is called the feature vector or pattern vector. H_i is the hypothesis that the pattern occurs from the ith of m pattern classes. The decision is the conclusion that H_i is true, with i specified. The feature extraction problem, or "characterization" problem, is to find a set of features suitable for use in the classification process. The features selected should be the most informative of the various properties or attributes of the situation or object. Feature selection is generally the most difficult and the most important process in designing a pattern recognition system.

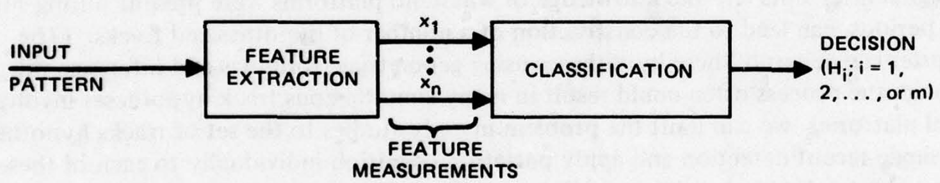

Figure 7. A pattern recognition system, shown as a division into two functions.

The optimum procedure (for minimizing the probability of misclassification) is to use a Bayes Classifier, which reduces to the maximum likelihood rule when the classes are equally likely in occurrence. The m decision functions

$$D_i(x) = p(x \mid H_i) \qquad i = 1, 2, \ldots, m$$

are calculated and the maximum $D_i(x)$ corresponds to the decision that H_i is true. Our difficulty with this procedure is that we must provide the conditional densities $p(x \mid H_i)$, and,

---

22. Ho YC and Agrawala AK, "On pattern classification algorithms – introduction and survey," Proc IEEE, vol 56, p 2101–2114, Dec 1968. Reprinted in Machine Recognition of Patterns, p 247–260, IEEE Press, 1977.
23. Fukunaga, K, Introduction to Statistical Pattern Recognition, Academic Press, 1972.
24. Patrick EA, Fundamentals of Pattern Recognition, Prentice-Hall, 1972.
25. Kanal L, "Patterns in pattern-recognition: 1968–1974," IEEE Trans on Information Theory, vol IT-20, no 6, p 697–722, Nov 1974. Reprinted in Machine Recognition of Patterns, p 1–26, IEEE Press, 1977.

even when these are known, the calculations can be prohibitively long. There are many applicable near-optimum and suboptimum pattern recognition classification procedures described in the literature that are simpler to implement than the optimum procedure. For the kinds of pattern recognition problems we expect to encounter in data fusion, the selection and implementation of a suitable decision procedure should be relatively simple compared to the feature extraction task.

Applications to Data Fusion Problems

First we will consider some examples of data fusion problems where pattern recognition might be applied.

Changes in course and speed. The observed track of a platform sometimes can be recognized by an operator as, for example, a change of station with respect to a formation guide. Other reasons for maneuvers are to evade a submarine or to avoid a known sensor. Also, a merchant likely will change his course to avoid bad weather while a warship might hide from sensor detection by following bad weather. Recognition of track patterns can help to distinguish between warships and other ships and can also indicate if the enemy knows of the presence of a submarine, a sensor, etc.

In the situation considered here, the tracks would be constructed by a computer using multisource inputs. Reported detections over a large area and over many hours, plus signal intercepts and the knowledge of where no platforms were present during surveillance periods, can lead to the construction of a number of hypothesized tracks. (The computer can generate these hypotheses using geometrical formulas and inference rules.) Although the process often could result in many simultaneous track hypotheses involving several platforms, we can limit the problem in early studies to the set of tracks hypothesized for a single recent detection and apply pattern recognition individually to each of these tracks. The track-type hypotheses $\{H_i\}$ would include possible kinds of maneuvers, and some of the measurements would relate to the presence of bad weather, a formation guide, a submarine, etc. If an uncertain track involves a maneuver that cannot be interpreted as an evasive course or to be related to weather or other causes of course changes, it is probably safe to reject that track hypothesis. If a single firm track is constructed from multisource data and it matches a maneuver pattern, additional information about that platform is gained.

Platform identification. The feature measurements used in the classification process would relate to characteristics or attributes of the platform. For example, feature measurements could include (1) indications of the ship structure and of its active sensors, as learned by passive sensors (signal intercepts, sonar informaton about ship noises, etc.), (2) indications of the ship structure learned from active sensors, such as radar and sonar indications of size or shape; (3) indications of behavior attributes such as maneuvers or erratic use of radar, as determined from active or passive sensors or other observations. One difficulty in applying pattern recognition to the problem of identifying a platform based on multisource data is that often only a subset of features will be observed or measured. In some cases, the absence of features in itself will be information; e.g., the lack of intercepted signals could imply intended electronic silence. The feature selection problem would be different from that in

most pattern recognition problems, since the feature measurements will come from various sources. Also, the pattern recognition process must allow for very contradictory evidence, instead of for the usual situation where measurements are noisy but none are entirely incorrect.

Situation recognition. The recognition of situations is the principal function of the fusion processes, while the two processes given as examples just above are steps in the direction of recognizing situations. Although we are considering here the application of pattern recognition to the recognition of situations, situation recognition is treated by some as a field in itself. The difference between pattern recognition and situation recognition is described by two Russian authors (ref. 26, p 70–71).

> "Situation recognition is a new branch of cybernetics; established terminology and voluminous literature are still lacking; individual publications are narrowly specialized in character. The most closely related field is pattern recognition, but there is a fundamental difference. First, a pattern is static and a situation is dynamic. Second, situation recognition always involves prediction, foresight, and extrapolation, which is usually not the case in pattern recognition. Third, pattern recognition presumes the existence of a classification system, and a basic finite alphabet of patterns established by training. When a new pattern is shown it is necessary to decide to which class it belongs (or to decice that it does not belong to any class). There is no a priori classification in situation recognition, since the number of possible situations is infinite, even though the results have a classification and a finite alphabet. Moreover, various situations may be similar and may even partially overlap in terms of the *initial* state and character of process. Expressed mathematically, many situations are *continuous* (i.e., such that a third, intermediate situation can always be found between two others), while many patterns are *never continuous*. This property of situations is a serious barrier to their recognition."

The authors continue by distinguishing three types of situations: simple, complex and degenerate. Probably, at best, we can hope only to recognize simple situations, using automatic techniques. They give this definition of simple situations.

> "Simple situations are those which are classified beforehand and, consequently, whose characteristics are known. The alphabet of simple situations is finite; it is assumed to be completely known to a commander and his staff, even though it is constantly being supplemented during accumulation of experience."

Three examples of simple situations given in this book are a tank attack, command and staff training and army inspection of independent activity.

---

26. Druzhinin VV and Kontorov DS, Concept, Algorithm, Decision, Moscow, 1972. (Translated and published under the auspices of the U.S. Air Force; Superintendent of Documents, U.S. Government Printing Office.)

The kinds of situations most amenable to automatic recognition could better be described as states. For example, consider a situation assessment problem where the state of combat readiness of a hostile task unit or group is to be determined. The categories might be (a) prepared for major conflict, (b) preparing for major conflict (moderate preparedness and building up), (c) staying at moderate readiness, (d) inadequate readiness but building up, and (e) staying at inadequate readiness. An infrequent but periodic run of a pattern recognition routine would use, as features, indicators generated from a data base of recent observations. As in the example for platform identification, we have feature measurements from a variety of sources, which is not the usual case in pattern recognition. Another difference from the usual kind of pattern recognition is that the feature selection process must be based on "prediction, foresight and extrapolation" (characteristics noted in the earlier quote), unless there have been recent major conflicts and much information about enemy readiness plus the associated evidence about that readiness. Also, we are dealing with continuous situations, although the fact that there is an inherent intent in each of the categories helps to justify our treating them as a finite alphabet of states. Still, the mechanics of this situation recognition process would be that of pattern recognition.

Sequential Methods

The use of a sequential decision procedure in the pattern recognition process is practical when the cost of taking feature measurements is significant or if the features are extracted sequentially in nature (refs. 22, 27, 28).

At each stage of a sequential decision process either a terminal decision is made or the decision to take an additional measurement. Ordering the features so that the most informative are used first will cause the terminal decision to be made earlier.

In the pattern recognition applications that have been considered here, we cannot specify beforehand a specific set of features and proceed to take measurements. Since the set of measured features can vary from one application to the next of the same pattern recognition process, and since updates in the measurements will sometimes occur during a single application, a sequential decision procedure seems highly appropriate. The procedure would first use data readily available, and in some cases would attempt to acquire new data if needed, by querying a remote data base or even by recommending an act of reconnaissance (an active form of fusion which we do not intend to consider in this project). Optionally, a tentative decision could be outputted when an early, tentative-decision bound is crossed, and the process would continue so long as profitable (until truncation) or until a small-error decision bound is crossed.

Pattern Recognition Versus Production Rules

In this section we have considered briefly the possible application of pattern recognition to the problem of platform identification. In Section III.B we discussed the application of production rules to the same problem, although the emphasis in that discussion was on a different kind of data. Augmenting a production system with new modular units of knowledge (e.g., If emission is type X and deception is unlikely then the class is

---

27. Fu KS, Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, 1968.
28. Fu KS, On Sequential Pattern Recognition Systems, in Methodologies of Pattern Recognition, Academic Press, 1969.

probably A or B) is relatively easy in a well-structured production system while a pattern recognition system generally would need redesigning. User confidence is another comparative advantage of the production system, because the rules employed by the system and the logic behind any decision are available to the user. On the other hand, the production system format is a clumsy structure for propagating weights while a pattern recognition system can efficiently use probability distributions or whatever information is available. Also, a pattern recognition system can be designed to cope with subtle differences in input data and to contain little redundancy.

For fusion problems such as this, where both pattern recognition and production rules appear to be applicable, comparisons need to be made. The appropriate system structure for each of the two approaches should be formulated and investigations made to determine which of the two is better for a particular application.

Interference with Other Techniques

Since a pattern recognition technique generally is applicable only to well defined and relatively static situations or pattern-classes, its use in data fusion most likely would occur as a specialized process embedded in a more general fusion process. (Recall the example of recognizing track patterns that involve maneuvers, evasive actions, and weather avoidance. These classifications would be needed in evaluating premises of certain production rules.) In such a case, it probably would not directly interface with data processing techniques (or with the data bases generated from natural language data and formatted text), but would begin with partially fused data. In some cases, the more general process would have to select the appropriate set of hypotheses $\{H_i\}$, and then initialize and trigger the pattern recognition routine.

IV. CONCLUSIONS

Several of the newer technology areas have been examined for their application to automatic fusion of multisource data. A review of some of the current work in the area of natural language processing (NLP) showed that the most applicable approaches to converting narrative data to conceptually structured data usually involve the use of "frames" of some kind, a frame being a data structure network designed to represent a situation. A text understanding technique involving "scripts" and "plans," special versions of a frame, is being developed at Yale University. This method is especially interesting because of its use of "theme" knowledge to determine the goal that underlies a plan, a process which can lead to a proper interpretation of the actions in a text. It is too early to determine whether or not this technique or others currently being developed elsewhere will be adequate for future use in an automated data fusion system, but the prospect does appear favorable.

Two aspects of NLP for data fusion that will present special problems are: (1) Ellipses – unformatted message text is often expressed in abridged and incomplete sentences, with words such as "a" and "the" missing; and (2) the data base of conceptually structured data obtained from narrative text will need continual updating as new narrative data arrives. Besides finding solutions to these two special problems, fusion of natural language (NL) data with other data will require finding appropriate ways of interfacing the processed NL data with the automated fusion processes that use them. It was pointed out in Section III.A that in several respects the NLP stage is not entirely separable from the

automatic fusion stage, but that some NLP is involved in fusion and some fusion is involved in NLP. The interfacing of NLP techniques with data fusion techniques will be a major concern in designing an automated data fusion system.

Fusion of many kinds of data ideally should result in comprehensible pictures or descriptions of situations (as complete as the data will support), with possible explanations of the goals and plans underlying the reported actions available to the fusion system user, along with any reasonable projections of future actions. Automated fusion will require the integration of many kinds of computerized processes, and it may be that a certain amount of human interaction and intervention will always be required. The concept of an integrated fusion system is still very nebulous, but a clear concept should evolve as we continue to look at specific techniques and consider how they must interact with other techniques. Our own investigation of techniques is limited to the newer technologies, but the problem of suitably interfacing these techniques with conventional and analytical techniques and with the human user must be carefully considered.

The particular fusion technique given the most attention during these initial investigations was the use of production rules to represent the knowledge and the chain of reasoning of a human operator or intelligence analyst. The attractive features of a production system are those that contribute to user confidence in the system and to user ease in modifying or expanding the system. Building the control mechanisms and the natural-like language interfaces that provide these features for the user should not be a prohibitively difficult task. Evaluating the premise sides of the rules based on mixtures of NL data and formatted data also should be possible, once a suitable method of NLP is sufficiently developed and adapted to this use. Probably the greatest problem will be to develop a production system organization that will support fast and efficient operation even when the number of rules is very large. Even if combined with advanced natural language understanding techniques, a production system will have to incorporate a tremendous amount of world knowledge, in the form of rules, in order for it to handle nonroutine situations, and there will be many nonroutine and new situations occurring which we would want a data fusion system to recognize. Another necessary complexity in a production system for this application is a weighting mechanism which will properly use estimates of certainty about the data and about the rules. Several approaches to weighting were described in Section III.B and the Appendix, and the investigation of weighting methods will continue through next year.

The application of pattern recognition methods to data was also examined. A few examples of applications were described in Section III.C, but none of these were in a problem form that a standard pattern recognition technique would nicely fit. Still, the general procedure of using measurements of features in a classification algorithm appears to have a few useful applications in data fusion, even though it cannot provide the user confidence and convenience attributes that a production system can. Investigations in this area should continue.

A brief look was given to the possibility of applying the theory of possibilities to data fusion. While no direct application was evident, it was found that fuzzy-set logic can be indirectly employed in other processes. An application mentioned in this report was the use of fuzzy-set computations in production system weighting. A fuzzy-set decision process is useful in pattern recognition when there are no precise boundaries between categories

and statistical independence cannot be assumed (ref. 29). Other possible applications of fuzzy sets are in the expression of effectiveness measures and in manipulations of real-world data (ref. 30). Further attention will be given to these and other possible applications.

---

29. Shimura, M, An Approach to Pattern Recognition and Associative Memories using Fuzzy Logic, in Fuzzy Sets and their Applications to Cognitive and Decision Processes, Academic Press, 1975.
30. Dockery JT, "The use of fuzzy sets in the analysis of military command," in Proc. 39th MORS, 1978.

REFERENCES

1. Minsky M, A Framework for Representing Knowledge, in the Psychology of Computer Vision, ed. PH Winston, p 211-277, McGraw-Hill, 1975.

2. Minsky M, Minsky's Frame System Theory, in Proceedings of the Conference on Theoretical Issues in NLP, p 104–116, Cambridge Mass, June 1975.

3. Winston PH, Artificial Intelligence, Addison-Wesley, 1977.

4. Hewitt C, Stereotypes as an ACTOR Approach Towards Solving the Problem of Procedural Attachment in FRAME Theories, in Proceedings of the Conference on Theoretical Issues in NLP, p 94–103, Cambridge Mass, June 1975.

5. Feldman J, Bad-Mouthing Frames, in Proceedings of the Conference on Theoretical Issues in NLP, Cambridge Mass, p 92-93, June 1975.

6. Schank RC, and Abelson RP, Scripts Plans and Knowledge, Advanced Papers for the Proceedings of the Fourth International Joint Conference on Artificial Intelligence, p 151-157, Tbilisi USSR, September 1975.

7. Schank RC, and Abelson RP, Scripts Plans Goals and Understanding; An Inquiry into Human Knowledge Structures, Lawrence Erlbaum Associates, 1977.

8. Schank RC, Identification of Conceptualizations Underlying Natural Language, in Computer Models of Thought and Language, ed. RC Schank and KM Colby, p 187-247, WH Freeman and Co, 1973.

9. Schank RC, The Primitive ACTs of Conceptual Dependency, in Proceedings of the Conference on Theoretical Issues in NLP, p 34-37, Cambridge Mass, June 1975.

10. Schank RC, Using Knowledge to Understand, in Proceedings of the Conference on Theoretical Issues in NLP, p 117-121, Cambridge Mass, June 1975.

11. Raphael B, The Thinking Computer, WH Freeman and Co, 1976.

12. Silva G, and Montgomery CA, Knowledge Representation for Automated Understanding of Natural Language Discourse: Computers and the Humanities, vol 11, p 223-234, Pergamon Press, 1978.

13. Charniak E, Organization and Inference in a Frame-Like System of Common Sense Knowledge, in Proceedings of the Conference on Theoretical Issues in NLP, p 42-51, Cambridge Mass, June 1975.

14. Electr Res Lab, Univ of Calif Berkeley, Unclassified Memorandum ERL-M77/61, Subject: PRUF-A Meaning Representation Language for Natural Languages, by Zadeh LA, 30 August 1977.

15. Zadeh LA, PRUF and Its Application to Inference from Fuzzy Propositions, vol 2, p 1359–1360 of Proceedings of New Orleans IEEE Conference on Decision and Control, IEEE publication, 1977.

16. Zadeh LA, A Theory of Approximate Reasoning (AR), Memorandum UCB/ERL M77/58, Electr Res Lab, Univ of Calif, Berkeley, 30 Aug 1977.

17. Davis R, BG Buchanan, and EH Shortliffe, Production Rules as a Representation for a Knowledge-Based Consultation System, Stanford AI Lab Memo AIM-266, Computer Science Dept Rept STAN-CS-75-519, Oct 1975.

18. Davis R and King J, An Overview of Production Systems, Stanford AI Lab Memo AIR-270, Computer Science Dept Rept STAN-CS-75-524, Oct 1975.

19. Duda RO, Hart PE, and Nilsson NJ, Subjective Bayesian Methods for Rule-Based Inference Systems, SRI-AI Center Tech Note 124, Jan 1976.

20. Hayes-Roth F, "Knowledge representation, organization, and control in large-scale pattern-based understanding systems," Conf Record, Joint Workshop on Pattern Recog and Artif Intel, p 66-73, 1976.

21. Shortliffe EH, Computer-Based Medical Consultations: MYCIN, American Elsevier, 1976.

22. Ho YC and Agrawala AK, "On pattern classification algorithms – introduction and survey," Proc IEEE, vol 56, p 2101–2114, Dec 1968. Reprinted in Machine Recognition of Patterns, p 247–260, IEEE Press, 1977.

23. Fukunaga K, Introduction to Statistical Pattern Recognition, Academic Press, 1972.

24. Patrick EA, Fundamentals of Pattern Recognition, Prentice-Hall, 1972.

25. Kanal L, "Patterns in pattern-recognition: 1968–1974," IEEE Trans on Information Theory, vol IT-20, no 6, p 697–722, Nov 1974. Reprinted in Machine Recognition of Patterns, p 1–26, IEEE Press, 1977.

26. Druzhinin VV and Kontorov DS, Concept, Algorithm, Decision, Moscow, 1972. (Translated and published under the auspices of the U.S. Air Force; Superintendent of Documents, U.S. Government Printing Office.)

27. Fu KS, Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, 1968.

28. Fu KS, On Sequential Pattern Recognition Systems, in Methodologies of Pattern Recognition, Academic Press, 1969.

29. Shimura M, An Approach to Pattern Recognition and Associative Memories using Fuzzy Logic, in Fuzzy Sets and their Applications to Cognitive and Decision Processes, Academic Press, 1975.

30. Dockery JT, "The use of fuzzy sets in the analysis of military command," in Proc. 39th MORS, 1978.

SUPPLEMENTARY BIBLIOGRAPHY

[The following Publications were not referenced directly, but they contributed to the ideas expressed in this document. Many additional publications were also helpful but are not listed here because of their restricted distribution.]

Abelson RP, The Structure of Belief Systems, in Computer Models of Thought and Language, WH Freeman, 1973.

Amarel S, "Computer-based modeling and interpretation in medicine and psychology; the Rutgers Research Resource," Federation Proc, vol 33, no 12, p 2341-2347, Dec 1974.

Biss K, Chien RT, Stahl F, and Weissman SJ, "Semantic modeling for deductive question answering," IEEE Trans on Computers, vol C-25, no 4, p 358-366, Apr 1976.

Chandrasekaran B, "Artificial intelligence – the past decade," in Advances in Computers, vol 13, ed by M Rubinoff and MC Yovits, New York: Academic Press, p 169-232, 1975.

Hart PE, Progress on a Computer-Based Consultant, SRI-AI Group – Technical Note 99, Jan 1975.

Hawkinson L, "The representation of concepts in OWL," Advance Papers of the Fourth Intern Joint Conf on Artif Intel, vol 1, p 107-114, Sep 1975.

Hendrix GG, "Expanding the utility of semantic networks through partitioning," Advance Papers of the Fourth Intern Joint Conf on Artif Intel, vol 1, p 115-121, Sep 1975.

Hodge DC, and HE Webb, Automatic Pattern Recognition in the United States, Part 1: DOD Supported Projects; Part 2: Projects Supported by Other Sources, U.S. Army Engineering Laboratory, Sep 1976.

Jackson Jr. PC, Introduction to Artificial Intelligence, Petrocelli books, 1974.

Jain R, "Decisionmaking in the presence of fuzziness and uncertainty," Proc Special Symp on Fuzzy Set Theory and Applic, vol 2 of Proc 1977 IEEE Conf on Decision and Control, p 1318-1323, 1977.

Kelly E, and Stone P, Computer Recognition of English Word Senses, vol 13, North-Holland Linguistic Series, 1975.

Lesser VR, Fennell RD, Erman LD, and Reddy DR, "Organization of the Hearsay II Speech Understanding System," IEEE Trans on Acoustics, Speech, and Signal Processing, vol ASSP-23, no 1, p 11-24, Feb 1975.

Maroy JP, "Natural language understanding as a pattern recognition problem," Proc Third Intern Joint Conf on Pattern Recognition, IEEE Publication, p 411–418, Nov 1976.

Palme J, "Making computers understand natural language," in Artificial Intelligence and Heuristic Programming, ed by NV Dindler, B Meltzer, New York: American Elsevier, p 199–244, 1971.

Raphael B, "The frame problem in problem-solving systems," in Artificial Intelligence and Heuristic Programming, ed by NV Findler and B Meltzer, New York: American Elsevier, p 159–169, 1971.

Reeker LH, "The computational study of language acquisition," in Advances in Computers, vol 15, edited by M Rubinoff and MC Yovits, New York: Academic Press, 1976.

Rieger C, "The commonsense algorithm as a basis for computer models of human memory, inference, belief and contextual language comprehension," Proc Conf on Theoretical Issues in NLP, Cambridge, Mass., p 180–195, 1975.

Rieger C, "Conceptual overlays: a mechanism for the interpretation of sentence meaning in context," Advance Papers of the Fourth Intern Joint Conf on Artif Intel, vol 1, p 143–150, Sep 1975.

Sacerdoti ED, "The nonlinear nature of plans," Advance Papers of the Fourth Intern Joint Conf on Artif Intel, vol 1, p 206–214, Sep 1975.

Sacerdoti ED, Language Access to Distributed Data with Error Recovery, SRI TN 140, Feb 1977.

Schank RC, "Finding the conceptual content and intention in an utterance in natural language conversation," Second Intern Joint Conf on Artif Intel, London, p 444–454, 1971.

Simmons RF, "Semantic networks: their computation and use for understanding English sentences," in Computer Models of Thought and Language, ed by RC Schank and KM Colby, San Francisco: WH Freeman, 1973.

Stearns SD, "On selecting features for pattern classifiers," Proc Third Intern Joint Conf on Pattern Recog, IEEE Publication, 71-75, Nov 1976.

Thompson FB, and Thompson BH, "Practical natural language processing: the REL system as prototype," in Advances in Computers, vol 13, ed by M Rubinoff and MC Yovits, New York: Academic Press, 1975.

Watanabe S, "Pattern recognition as an inductive process," in Methodologies of Pattern Recognition, ed by S Watanabe, New York: Academic Press, 1969.

Winograd T, "A procedural model of language understanding," in Computer Models of Thought and Language, ed by RC Schank and KM Colby, San Francisco: WH Freeman, 1973.

Woods WA, "Some methodological issues in NL understanding research," Proc Conf on Theoretical Issues in NLP, Cambridge, Mass., p 134–139, 1975.

Zadeh LA, "Calculus of fuzzy restrictions," in Fuzzy Sets and their Applications to Cognitive and Decision Processes, ed by L Zadeh, K Fu, K Tanaka, and M Shimura, New York: Academic Press, 1975.

APPENDIX

Two methods of propagating weights in a production system are summarized below. They are Bayesian Weighting and Belief Measures.

Bayesian Weighting

The results given here are essentially a summary of some derived by Duda, et al., in reference 19, although the notation and the order of presentation have been changed.

Single rule case. Consider the rule $F \rightarrow C$. Let W_o denote the prior odds on C

$$W_o \triangleq P(C)/(1 - P(C)) \ ,$$

and let the "strength" of the rule be

$$A \triangleq P(F|C)/P(F|\overline{C})$$

for F true and

$$\overline{A} \triangleq P(\overline{F}|C)/P(\overline{F}|\overline{C})$$

for F false. Let E denote evidence about F, and let $D = P(F|E)$ denote the *certainty factor*. The updating formula for finding the posterior odds

$$W(E) \triangleq P(C|E)/P(\overline{C}|E)$$

is

$$W(E) = \frac{DP(C|F) + (1 - D)P(C|\overline{F})}{DP(\overline{C}|F) + (1 - D)P(\overline{C}|\overline{F})} \tag{A-1}$$

where

$$P(C|F) = AW_o/(1 - AW_o)$$

and

$$P(C|\overline{F}) = \overline{A}W_o/(1 - \overline{A}W_o) \ .$$

For $P(F|E) = 1$, (A-1) gives $W(E) = W(F) = AW_o$, and for $P(\overline{F}|E) = 1$, it gives $W(E) = W(\overline{F}) = \overline{A}W_o$.

Weighting at AND nodes. When the left side of the rule is a conjunction

F_1 and F_2 and ... and $F_n \rightarrow C$

then let F denote the event that all F_i are true and E the evidence $E_1, E_2, \ldots, E_n$, and use (A-1) directly. If the F_i are independent (conditionally on H and $\bar{H}$) and also the E_i, then

$$D = \sum_{i=1}^{n} P(F_i | E_i) \tag{A-2}$$

Weighting at OR nodes. When we have several rules

$F_1 \rightarrow C, \qquad F_2 \rightarrow C, \ \ldots, \qquad F_n \rightarrow C$

all concerning the same hypothesized conclusion C, and we can as above assume independent evidence, then the updating formula for finding the posterior odds

$$W(E_1, \ldots, E_n) \triangleq \frac{P(H | E_1, \ldots, E_n)}{P(\bar{H} | E_1, \ldots, E_n)}$$

is

$$W(E_1, \ldots, E_n) = W_o \sum_{i=1}^{n} \frac{W(E_i)}{W_o} \tag{A-3}$$

Inconsistencies. In practice, there are problems encountered in using Bayesian updating when dealing with collections of subjective inference rules. It is explained in reference 19 that these Bayesian results are valid if the prior odds W_o and the strengths A and $\bar{A}$ are specified consistently, but that they are virtually certain to be inconsistent. Several measures that can be taken to correct the effects of priors that are inconsistent with inference rules are summarized in reference 19.

Belief Measures

Section III.B discussed a method of weighting used in the MYCIN system, a method which could be modified for use in production systems for platform identification and other data fusion problems. The method is summarized below, based on a description by Shortliffe in Chapter 4 of reference 21. The notation and terminology has been changed in order to be more consistent with that used elsewhere in this report.

Consider first the simple production rule $F \rightarrow C$. The measure of increased belief in C, based on F, is defined as

$$mb[C,F] = \begin{cases} 1 & \text{if } P(C) = 1 \\ \dfrac{\max[P(C|F), P(C)] - P(C)}{1 - P(C)} & \text{otherwise.} \end{cases} \tag{A-4}$$

The measure of increased disbelief in C, based on F, is defined as

$$md[C,F] = \begin{cases} 1 & \text{if } P(C) = 0 \\ \dfrac{\min[P(C|F), P(C)] - P(C)}{-P(C)} & \text{otherwise .} \end{cases} \tag{A-5}$$

These two measures are graphed in figures A-1 and A-2. The certainty factor of the rule is defined as

$$cf[C,F] = mb[C,F] - md[C,F] \tag{A-6}$$

which can also be written

$$cf[C,F] = \begin{cases} 1 & \text{if } P(C) = 1 \\ \dfrac{P(C|F) - P(C)}{1 - P(C)} & \text{if } P(C|F) > P(C) \\ 0 & \text{if } P(C|F) = P(C) \neq 0 \text{ or } 1 \\ \dfrac{P(C|F) - P(C)}{-P(C)} & \text{if } P(C|F) < P(C) \\ -1 & \text{if } P(C) = 0 \end{cases} \tag{A-7}$$

The certainty factor of F, based on evidence E, is defined in the same manner. The measures of increased belief or disbelief in C, based on evidence E, are approximated by the formulas

$$mb[C,E] = mb[C,F] \cdot \max(0, cf[F,E]) \tag{A-8}$$

$$md[C,E] = md[C,F] \cdot \max(0, cf[F,E]) \ . \tag{A-9}$$

The certainty factor of C, based on E, is given by the definition $cf = mb - md$. Substituting [F,E] for [C,F] in (A-7) to find max(0, cf[F,E]), we obtain the approximation

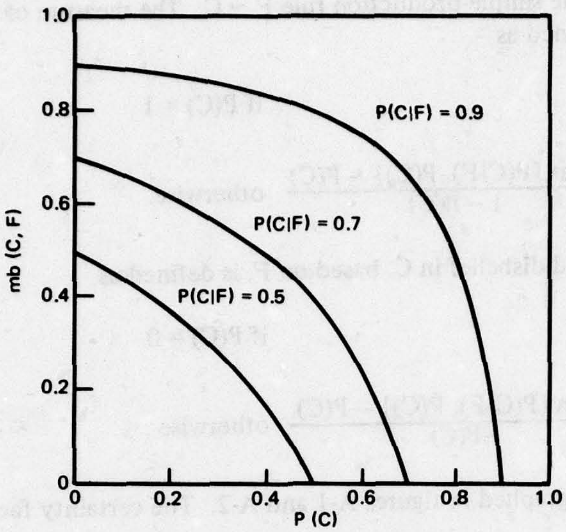

Figure A-1. The measure of increased belief in C, based on F, for P(C|F) > P(C). The measure mb[C,F] is zero for P(C|F) ⩽ P(C) ≠ 1 and is unity for P(C) = 1. If P(C) - 1 or 0, then P(C|F) = P(C).

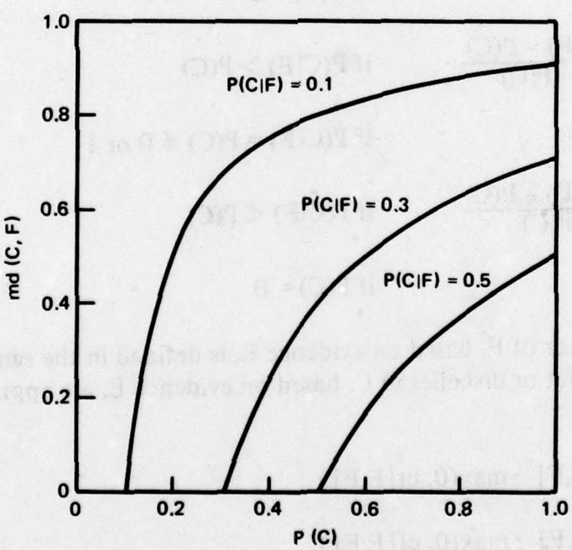

Figure A-2. The measure of increased disbelief in C, based on F, for P(C|F) < P(C). The measure md[C,F] is zero for P(C|F) ⩾ P(C) ≠ 0 and is unity for P(C) = 0.

$$
cf[C,E] = \begin{cases} cf[C,F] & \text{if } P(F) = 1 \\ cf[C,F] \cdot \dfrac{P(F|E) - P(F)}{1 - P(F)} & \text{if } P(F|E) > P(F) \\ 0 & \text{otherwise .} \end{cases} \tag{A-10}
$$

If F is required for C to be true (i.e., $\bar{F} \rightarrow \bar{C}$), and if $P(F|E) < P(F)$, then one would desire that cf[C,E] be negative. Note, however, that (A-10) gives $cf[C,E] = 0$ for $P(F|E) \leqslant P(F)$. This property is not a problem for any of the sample rules given for platform identification in Section III.B, but is unsatisfactory for rules of the type, "If the ship has n screws and m blades then it is class x." If evidence E indicates that a ship does not have n screws or m blades (assuming that measurements of screw propeller characteristics can be obtained), then the conclusion that it is class x should have a negative certainty factor.

Incrementally acquired evidence. Consider the rule

$$
F_1 \,\&\, F_2 \,\&\, \ldots \,\&\, F_n \rightarrow C \ .
$$

If the certainty factor $cf[C,F_1 \,\&\ldots\&\, F_n]$ of this rule is not specified by the expert but the individual certainty factors $cf[C,F_i]$ are, then the following approximation technique can be used.

The measure of the increased belief in C, based on $F_1 \,\&\ldots\&\, F_n$, is approximated by using the formula

$$
mb[C,F_1\&\ldots\&F_n] \begin{cases} 0 & \text{if } md[C,F_1 \,\&\ldots\&\, F_n] = 1 \\ mb[C,F_1\&\ldots\&F_{n-1}] + mb[C,F_n] \\ \quad \cdot (1 - mb[C,F_1\&\ldots\&F_{n-1}]) \\ \quad \text{otherwise .} \end{cases} \tag{A-11}
$$

Note that this approximation gives $mb[C,F_1 \,\&\ldots\&\, F_n] = 1$ if $mb[C,F_i] = 1$ for any i. In this respect, the formula treats the node as an OR node instead of an AND node. The measure of increased disbelief is approximated in the same manner, and the certainty factor of C, based on $F_1 \,\&\ldots\&\, F_n$ is given by the definition $cf = mb - md$. Using (A-8) and (A-9), we have

$$
cf[C,E] = \begin{cases} cf[C,F_1\&\ldots\&F_n] & \text{if } P(F_1\&\ldots\&F_n) = 1 \\ cf[C,F_1\&\ldots\&F_n] \cdot \max(0, cf[F_1\&\ldots\&F_n,E]) & \\ & \text{if } P(F_1\&\ldots\&F_n|E) \\ & \quad > P(F_1\&\ldots\&F_n) \\ 0 & \text{otherwise .} \end{cases} \tag{A-12}
$$

The certainty factor $cf[F_1 \& \ldots \& F_n, E]$, where E denotes collectively the evidence for all F_1, can be approximated by using the definition mf = mb − md with the following formulas for conjunctions of hypotheses.

$$mb[F_1 \& \ldots \& F_n, E] = \min(mb[F_1,E], \ldots, mb[F_n,E]) \quad \text{(A-13)}$$

$$md[F_1 \& \ldots \& F_n, E] = \max(md[F_1,E], \ldots, md[F_n,E]) \quad \text{(A-14)}$$

Reference 21 also gives approximations for disjunctions of hypotheses. For determining the certainty factor (cf = mb − md) of $F_1 \ldots F_n$, based on E, these formulas would be

$$mb[F_1 V \ldots VF_n, E] = \max(mb[F_1,E], \ldots, mb[F_n,E]) \quad \text{(A-15)}$$

$$md[F_1 V \ldots VF_n, E] = \min(md[F_1,E], \ldots, md[F_n,E]) \quad \text{(A-16)}$$

The formulas for conjunctions and disjunctions of hypotheses, when used this way to estimate the certainty of the combined F_i's, distinguish between the AND combination $F_1 \& \ldots \& F_n$ and the OR combination $F_1 V \ldots VF_n$, while the combining formula for the certainty factor of C treats the node defined by $F_1 \& \ldots \& F_n \rightarrow C$ as something between an OR node and an AND node. No formulas are given in reference 21 for approximating $cf[C, F_1 V \ldots VF_n]$ when given only the factors $cf[C, F_i]$, $i = 1, \ldots, n$.

Shortliffe shows that the formulas for estimating $cf[C, F_1 \& \ldots \& F_n]$ have many desirable properties but also that they do not apply to some situations. For example, they do not work in a situation where $F_1 \& \bar{F}_2 \rightarrow C$ and $\bar{F}_1 \& F_2 \rightarrow C$, but $F_1 F_2 \rightarrow \bar{C}$. Another situation is when F_1 implies F_2. Also, the method becomes unworkable for applications in which a large number of observations must be grouped in the premise of a single rule, i.e., when n is large.

Conclusions. This approach to defining and propagating weights in a production system is not entirely suitable for a data fusion application such as platform identification. However, if the method were expanded to check for special cases (e.g., whether $\bar{F}_1 \rightarrow \bar{C}$ when $F_1 \rightarrow C$, or whether $F_1 \rightarrow F_2$, or whether $F_1 \& F_2 \rightarrow \bar{C}$ or C or neither when $F_1 \bar{F}_2 \rightarrow C$ and $\bar{F}_1 F_2 \rightarrow C$) and to use modified formulas in these cases, the results might be very good.